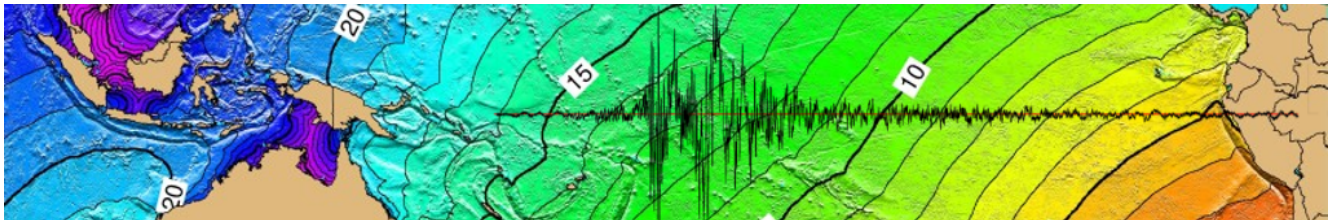


Minichallenge: Erdbebenmonitor



1 Ausgangslage

Du arbeitest für eine Rückversicherung, die unter anderem auch mit Naturkatastrophen verbundenen Risiken arbeitet. Für dein Team sollst du einen Newsfeed mit der aktuellen Erdbebensituation anbieten, der kontinuierlich aktualisiert wird und die Erdbeben in den letzten 24 Stunden anzeigt. Du greifst dazu auf die globalen Erdbebendaten von US Geological Surveys zu.

Konkret möchtest du über die JSON API auf die Daten zugreifen, die Erdbebeninformationen über die letzte Stunde, den letzten Tag, die letzte Woche und den letzten Monat anbietet. Die Links zu den einzelnen Datensätzen und eine Beschreibung der einzelnen Schlüsselwörter ist hier verfügbar:

<https://earthquake.usgs.gov/earthquakes/feed/v1.0/geojson.php>

Du schätzt den Aufwand für einen Proof-of-Concept auf einen (aufgerundeten) Tag und machst dich sofort an die Arbeit.

Zusatzinfo Die schwachen Erdbeben in der Schweiz werden davon selten erfasst, wenn dich diese auch interessieren, dann erhältst du dazu mehr Informationen auf der Seite des Swiss Seismological Service.

2 Aufgabenstellung

1. Daten einlesen

- a) Lade die Daten herunter und lese sie in eine geeignete Datenstruktur ein.

Hinweis: Dazu gibt es viele Varianten. Findest du die kürzeste?

- b) Exploriere nun die Struktur des JSON-Objekts, auch mithilfe der Beschreibungen der einzelnen Schlüsselwörter. Wie ist die Hierarchie aufgebaut?

2. Daten aufbereiten

- a) Versuche das JSON-Objekt mit geeigneten Manipulationen auf ein Data Frame 'flachzudrücken'. Wie kannst du passend die verschiedenen hierarchischen Levels vereinen?
- b) Entferne Spalten, die du für deine Problemstellung als unnötig empfindest.
- c) Konvertiere den Zeitstempel (Sekunden seit 1970, die sogenannte *Unix Epoche*) in ein Datumsformat und nenne diese neue Spalte 'date'.
- d) Falls noch nicht in a) geschehen: Stelle sicher, dass du die Koordinaten in Form von Spalten 'lon', 'lat' und 'depth' zur Verfügung hast. Sind Longitude und Latitude immer im zulässigen Bereich?

3. Explorative Analyse

- a) Wie ist die Verteilung der Erdbebenstärken (Magnituden) im Datenset?
- b) Wie sind die Tiefen der Epizentren verteilt?
- c) Sind Tsunamis vorgekommen? Wenn ja wo?
- d) Visualisiere alle Erdbeben geeignet auf einer Weltkarte. Bringe auch die Magnitude in die Visualisierung. Falls erwünscht kannst du die Karte auch interaktiv machen und mit Detailinformationen verlinken (Hier eignet sich *Plotly Express* gut).

4. Erdbeben-Newsfeed

- a) Erstelle nun eine Pipeline, die dir die Erdbebendaten herunterlädt, einliest und gemäss den vorherigen Aufgaben präprozessiert. Als Endprodukt soll ein Data Frame wie am Ende von Aufgabe 2 zur Verfügung stehen. Die Funktion soll als Argument die URL des JSON bekommen, dann kannst du deine Funktion auch mit Daten des letzten Monats laufen lassen ohne die Pipeline umzuschreiben. Wie lang braucht deine Pipeline, um das Data Frame zu ziehen und zu präprozessieren?
- b) Falls du noch Zeit hast, kannst du nun mit 'dash' (Python) oder 'flexdashboard' (R) ein Dashboard erstellen, dass dir Erdbeben mit Magnitude grösser 5 auf einer Weltkarte darstellt und regelmässig aktualisiert.